Optimization Algorithms
(ACM 41030)

Dr Lennon Ó Náraigh

Exercises #5

1. Does the OP
$$\min f(\boldsymbol{x}) = (y+100)^2 + \tfrac{1}{100}x^2$$
subject to $y - \cos x \geq 0$ have a finite or infinite number of local solutions? Use the KKT conditions to justify your answer.

Let
$$\mathcal{L}(\boldsymbol{x},\lambda) = (y+100)^2 + \tfrac{1}{100}x^2 - \lambda\left(y - \cos x\right).$$
We solve $\nabla_x \mathcal{L} = 0$. We have:
$$\frac{\partial \mathcal{L}}{\partial x} = \tfrac{1}{50}x - \lambda \sin(x),$$
$$\frac{\partial \mathcal{L}}{\partial y} = 2(y+100) - \lambda.$$

The following are the KTT conditions for the problem:
$$\begin{cases} \tfrac{1}{50}x - \lambda\sin(x) = 0,\ 2(y+100) - \lambda = 0 \\ y - \cos(x) \geq 0, \\ \lambda \geq 0, \\ \text{No Equality Constraints,} \\ \lambda(y - \cos(x)) = 0. \end{cases}$$

We look at two cases.

Case 1: We take $\lambda = 0$, hence $x = 0$ and $y = -100$. This is the global minimum, it is not feasible, so Case 1 is ruled out.

Case 2: We take $\lambda \neq 0$. So $y = \cos(x)$. We have:
$$\begin{aligned} \tfrac{1}{50}x &= \lambda\sin(x), \\ 2(y+100) &= \lambda. \end{aligned}$$

Divide these equations one by the other:

$$\frac{\frac{1}{50}x}{2[100+\cos(x)]} = \sin(x). \tag{1}$$

Hence, $x = 100[100+\cos(x)]\sin(x)$. This is a root-finding problem. It can be solved graphically by looking at the curves $y_1(x) = x$ and $y_2(x) = 100[100 + \cos(x)]\sin(x)$. The points of intersection of the two curves, $y_1(x) = y_2(x)$ give the roots.

Notice that the curve $y_1(x)$ is unbounded while the curve $y_2(x)$ is bounded by $\pm 100 \times 101$. Thus, there will only be finitely many points of intersection where $y_1(x) = y_2(x)$ and hence, only **finitely many** roots.

Another way to look at this problem is to start with the result that the constraint is active, such that $y = \cos(x)$. Then, the cost function can be re-parametrized as:

$$f(x, y = \cos(x)) = [\cos(x) + 100]^2 + \tfrac{1}{100}x^2 = \widetilde{f}(x).$$

The critical points are at $\mathrm{d}\widetilde{f}(x)/\mathrm{d}x = 0$, hence

$$2\,[\cos(x) + 100]\sin(x) = \tfrac{1}{50}x,$$

which is exactly Equation (1) again. Furthermore, for large x, $\widetilde{f}(x) \sim x^2/100$, and $\mathrm{d}\widetilde{f}/\mathrm{d}x \sim x/50$. Thus, $\mathrm{d}\widetilde{f}/\mathrm{d}x \neq 0$ for x sufficiently large. Hence, the roots of Equation (1) must be contained in an interval $(-R, R)$ and as the roots can only be discrete, there are only finitely many of them.

2. Let $\boldsymbol{v} : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth vector function, and consider the unconstrained OP

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}),$$

where

$$f(\boldsymbol{x}) = \max_{i\in\{1,2,\cdots,m\}} v_i(\boldsymbol{x}).$$

Reformulate this (generally non-smooth problem) as a smooth constrained problem.

We look at an example first:

$$\boldsymbol{x}_* = \arg \min_{(x,y)\in\mathbb{R}^2} \Big\{ \max\left[v_1(x,y), v_2(x,y)\right] \Big\}. \tag{2}$$

We look at the surfaces $z = v_1(x,y)$ and $z = v_2(x,y)$ in $\mathbb{R}^3$, shown in Figure 1. The figure suggests that the OP problem (2) can be recast as:

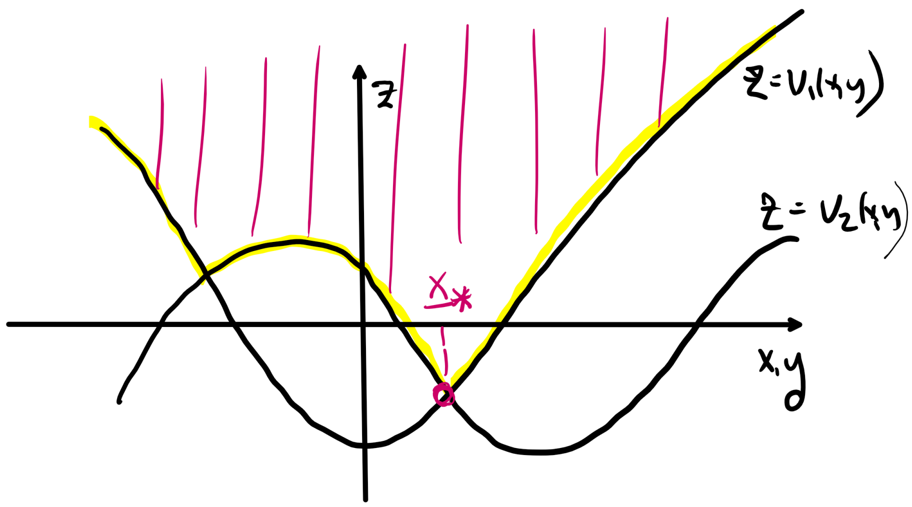


Figure 1:

$$\min_{(x,y,z)\in\Omega} z,$$

where

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 | z \geq v_1(x,y),\ z \geq v_2(x,y)\}.$$

This suggests the general reformulation of the OP in the original question is to introduce

$$\widetilde{\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{x} \\ z \end{pmatrix} \in \mathbb{R}^{n+1},$$

and we further introduce:

$$\Omega = \{\widetilde{\boldsymbol{x}} \in \mathbb{R}^{n+1} | z \geq v_1(\boldsymbol{x}), \cdots, z \geq v_m(\boldsymbol{x})\}.$$

We then consider the reformulated OP

$$\min_{\tilde{\boldsymbol{x}} \in \Omega} z.$$

3. Can you perform a smooth reformulation of the previous question when f is defined by:

$$f(\boldsymbol{x}) = \min_{i \in \{1,2,\cdots,m\}} v_i(\boldsymbol{x}).$$

Why or why not?

No. Inspect Figure 1 again. The min

$$f(\boldsymbol{x}) = \min_{i \in \{1,2,\cdots,m\}} v_i(\boldsymbol{x}).$$

is possibly unbounded.

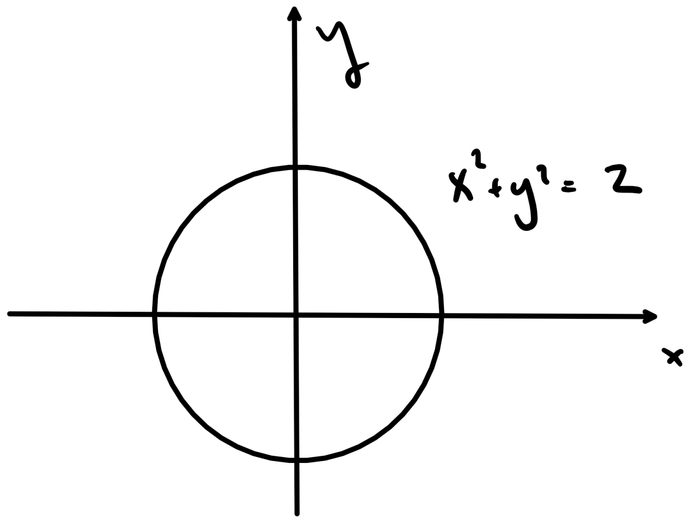


Figure 2:

4. Consider the OP

$$\min(x+y), \qquad \text{subject to } 2-x^2-y^2=0.$$

Specify two feasible sequences that approach the **maximizing** point $(1,1)^T$ and show that neither sequence is a decreasing sequence for f.

Let $\boldsymbol{x}_* = (1,1)^T = \sqrt{2}(\cos\pi/4, \sin\pi/4)^T$. We have:

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^2 | x^2+y^2=2\},$$

see Figure 2.

Consider the path

$$\boldsymbol{x}(\phi) = \sqrt{2}\left(\cos(\pi/4+\phi), \sin(\pi/4+\phi)\right)^T.$$

Introduce

$$\begin{aligned} \widetilde{f}(\phi) &= f(\boldsymbol{x}(\phi)), \\ &= \sqrt{2}\left[\cos(\pi/4+\phi)+\sin(\pi/4+\phi)\right] \end{aligned}$$

We have:

$$\widetilde{f}(\phi=0)=2, \qquad \left.\frac{\mathrm{d}\widetilde{f}}{\mathrm{d}\phi}\right|_{\phi=0}=0, \qquad \left.\frac{\mathrm{d}^2\widetilde{f}}{\mathrm{d}\phi^2}\right|_{\phi=0}=-2.$$

Hence,

$$\begin{aligned} \widetilde{f}(\phi) &= \widetilde{f}(0) + \left.\frac{\mathrm{d}\widetilde{f}}{\mathrm{d}\phi}\right|_{\phi=0}\phi + \tfrac{1}{2}\left.\frac{\mathrm{d}^2\widetilde{f}}{\mathrm{d}\phi^2}\right|_{\phi=0}\phi^2 + \cdots, \\ &= 2-\phi^2+\text{H.O.T.} \end{aligned}$$

We look at feasible sequences:

$$\boldsymbol{x}_n = \boldsymbol{x}(\phi_n), \qquad n \in \{1, 2, \cdots\},$$

where $\phi_n = 1/n$ or $\phi_n = -1/n$, such that $\boldsymbol{x}_n \in \Omega$, for all $n \in \{1, 2, \cdots\}$ and such that $\boldsymbol{x}_n \to \boldsymbol{x}_*$ as $n \to \infty$. Hence,

$$\begin{aligned} f(\boldsymbol{x}_n) &= \widetilde{f}(\phi_n), \\ &\overset{n\to\infty}{=} 2 - (\pm 1/n)^2. \end{aligned}$$

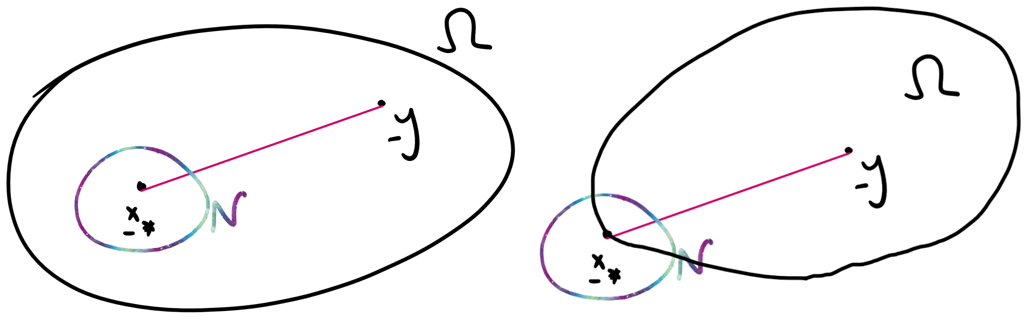
Thus, as $n \to \infty$,

$$\begin{aligned} f(\boldsymbol{x}_{n+1}) - f(\boldsymbol{x}_n) &= -\frac{1}{(n+1)^2} + \frac{1}{n^2}, \\ &= \frac{-n^2 + (n+1)^2}{n^2(n+1)^2}, \\ &= \frac{2n+1}{n^2(n+1)^2} > 0, \end{aligned}$$

and

$$f(\boldsymbol{x}_{n+1}) \geq f(\boldsymbol{x}_n), \qquad n \to \infty.$$

Hence finally, we see that $f(\boldsymbol{x}_n)$ is non-decreasing along the sequences $\boldsymbol{x}_n$, and $f(\boldsymbol{x}_n) \to f(\boldsymbol{x}_*)$ as $n \to \infty$.



(a) Option 1 (b) Option 2

Figure 3:

5. If f is convex and the feasible region Ω is convex, show that local solutions of the OP

$$\boldsymbol{x}_* = \arg\min_{\boldsymbol{x}\in\Omega} f(\boldsymbol{x})$$

are also global solutions.

Hint: Review Theorem 2.8 in the class notes.

Suppose that $\boldsymbol{x}_*$ is a local but not a global minimizer. Then we can find a point $\boldsymbol{y} \in \Omega$ such that $f(\boldsymbol{y}) < f(\boldsymbol{x}_*)$. By convexity of Ω, the line segment

$$L = \{\boldsymbol{x}(\alpha) | \boldsymbol{x}(\alpha) = \alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x},\ \alpha \in [0,1]\},$$

lies entirely in Ω. By convexity of f, we also have:

$$f(\boldsymbol{x}(\alpha)) \le \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}_*).$$

Refer to Figure 3. Any neighbourhood $\mathcal{N}$ around $\boldsymbol{x}_*$ contains a piece of the line segment L, so there exists an $\boldsymbol{x} \in \Omega$ such that $f(\boldsymbol{x}) < f(\boldsymbol{x}_*)$, which is a contradiction. Hence, $\boldsymbol{x}_*$ is the global minimizer.